WPS5208

Policy Research Working Paper 5208

Climate Cost Uncertainty, Retrofit Cost Uncertainty, and Infrastructure Closedown

A Framework for Analysis

Jon Strand
Sebastian Miller

The World Bank
Development Research Group
Environment and Energy Team
February 2010

Abstract

Large and energy-intensive infrastructure investments with long life times have substantial implications for climate policy. This study focuses on options to scale down energy consumption and carbon emissions now and in the future, and on the costs of doing so. Two ways carbon emissions can be reduced post-investment include retrofitting the infrastructure, or closing it down. Generally, the presence of bulky infrastructure investments makes it more costly to reduce emissions later. Moreover, when expected energy and environmental costs are continually rising, inherent biases in the selection processes for infrastructure investments lead to excessive energy intensity in such investments. Thus great care must be taken when choosing the energy intensity of the infrastructure at the time of investment. Simulations indicate that optimally exercising the retrofit option, when it is available, reduces ex ante expected energy consumption relative to the no-option case. Total energy plus retrofit costs can also be substantially reduced, the more so the larger is ex ante cost uncertainty. However, the availability of the retrofit option also leads to a more energy intensive initial infrastructure choice; this offsets some, but usually not all, of the gains from options for subsequent retrofitting.

This paper—a product of the Environment and Energy Team, Development Research Group—is part of a larger effort in the department to study analytical aspects of climate change. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at jstrand1@worldbank.org.

Climate Cost Uncertainty, Retrofit Cost Uncertainty, and Infrastructure Closedown: A Framework for Analysis

By

Jon Strand
The World Bank
Development Research Group
Environment and Energy Team

and

Sebastian Miller
Inter-American Development Bank
Southern Cone Department

Jstrand1@worldbank.org

1. Introduction

This paper presents a simple analytical model of infrastructure choice, and simulations to illustrate properties of the model framework. An important starting point is that infrastructure investments, sunk at an initial time of decision, "tie up" energy consumption for a long future period. In our model time is divided into two discrete periods: an initial period 1, "the present"; and a following period 2 ("the future"), which may be much longer. For each of the two periods, all relevant variables are assumed to be commonly known at the start of the period. Several key cost variables in period 2, including energy and environmental costs and certain technological costs (so-called "retrofit" costs, to be explained below), are unknown in period 1, but their ex ante statistical distributions for period 2 are assumed to be known when the initial infrastructure investment is made in period 1.[1] We assume that the infrastructure persists throughout both periods, but may be (deliberately, and at no additional cost) shut down at the start of period 2. The energy consumption tied to the infrastructure is by assumption based on fossil fuels, at least initially.

The focus of our analysis is that many types of infrastructure lead to considerable climate policy "inertia", in establishing levels of fossil-fuel consumption that may be difficult to reduce later. As increasingly recognized in the literature, including Ha-Duong et al (1996), Wigley (1996), Ha-Duong (1998), Lecocq et al (1998), and Shalizi and Lecocq (2009), the presence of such an established infrastructure may form a major *ex post* obstacle to effective mitigation policy, for a long future period, possibly 50-100 years or more. This is the case regardless of whether the initial infrastructure investment is "optimal" (in an ex ante sense), or not. The particular problem with infrastructure investment in this context is that costs of mitigation or abatement, to reduce emissions to desirable levels, may be very high ex post after the infrastructure has been established.

Below we discuss reasons why the infrastructure investment may be suboptimal, and then typically biased in the direction of too high fossil energy consumption and carbon emissions. The potential reasons are many and include systematic under-valuation of future energy costs; failures to incorporate true (current and future) social carbon emissions costs; and excessive discounting.

In our analysis we use a very simple model of infrastructure investment, where we introduce two potential mechanisms by which the fossil-fuel consumption can be modified "ex post" (in period 2). The first is to "retrofit" the infrastructure in period 2, at a cost. After a "retrofit", we assume, infrastructure operation causes no emissions of greenhouse gases (GHGs) from then on.

A "retrofit" can be interpreted in several ways. First, infrastructure may be operated without any use of fossil fuels from then on. This is particularly relevant when fossil fuels are replaced by alternative (non-fossil) energy sources, and the use of these sources reduces or eliminates the emissions of GHGs from the "normal" operation of the infrastructure. But this interpretation is also relevant when applied to cases where the overall energy demand of the

[1] This of course is a simplification at least relative to some presentations where the distribution of future energy and environmental cost is assumed to be unknown at an initial stage; this can lead to some serious problems of inference as argued e g by Weitzman (2009). Here, we may use a standard Bayesian approach to justify our position; namely, as the "best initial assessment" of the distribution given our current knowledge. See also Geweke (2001) and Schuster (2004) for formal justifications of this approach.

infrastructure is simply reduced. In our stylized model, emissions are then assumed to be eliminated completely. In either interpretation, the existence of a potential retrofit option in no way implies that retrofit is necessarily an economically optimal choice; our model includes cases in which exercising the option of retrofit would be prohibitively costly.

An alternative interpretation of “retrofit” is that the consumption of fossil fuels in operating the infrastructure is unchanged, but that the carbon is removed from these fuels (through carbon capture and storage, CCS, or similar technologies or processes). This will however not be our main interpretation in the following. It can be a somewhat problematic interpretation in our model, in particular when the post-retrofit operating period T is a variable, since the retrofit cost here is “periodized” and assumed constant “per time unit” within period 2. When T is variable, retrofit costs with a given distribution G will then correspond to a variable total retrofit cost (as it would be proportional to a variable T). When the retrofit cost in fact represents a given sunk cost initially in period 2, the G function will need to be amended when T changes.

Both “retrofit” costs and basic energy/climate cost are assumed to be uncertain from the point of view of period 1, but with initially known joint distribution. In our basic presentation and simulations, the two costs are assumed to be uncorrelated; this is however often unrealistic and we comment later on implications of positively correlated prices. As for the degree of uncertainty, retrofit costs could easily be more uncertain than energy and climate costs, since future (period 2) retrofit technologies are mostly unknown at the start of period 1 when initial infrastructure investments must be made.

The second way to avoid energy consumption related to existing infrastructure in period 2 is to simply abandon it, “close it down”. This is wasteful in the sense that the initial infrastructure investment was costly to establish, and this investment is then made worthless. It is a painful option, but can still be attractive and rational ex post, in states where energy and/or retrofit costs of continued operation both turn out to be very high at the same time. The required condition is that the lower of these costs is higher than the utility value of continued operation. The abandoned infrastructure will then, presumably, be replaced by alternative, less energy-intensive, infrastructure in period 2. Our model however does not specify any particular replacement alternative. In effect, our closedown alternative represents a “benchmark” case with zero emissions and energy consumption.[2]

We characterize ex ante strategies for establishing energy and emissions intensity associated with the initial infrastructure investment; ex post strategies for retrofitting and operating the infrastructure at a “later” stage (in period 2); and interactions between these strategies. An important issue is to study optimal infrastructure investment; both characterization, and (even more important from a climate policy standpoint) factors behind inefficient (too energy and emissions intensive) infrastructures. Another key issue is whether, and to what extent, an initially high energy intensity level can be modified in later periods through retrofit or closedown, in cases where energy and environmental costs are high.

“Optimal” infrastructure choice is defined for given current prices and distributions of future prices. “Optimality” can be established either for a private agent making the infrastructure

[2] One possible interpretation of this case is that energy consumption and emissions in the closedown alternative serve as a reference “zero” point, relative to the “business-as-usual” and retrofit alternatives.

decisions, or for a social planner; which if these will be invoked in the following will depend on the context. The decision-making agent could be private, but in most cases a public-sector entity (a local or national government). A social planner, taking a national, regional or global perspective, will tend to incorporate prices, costs, discount rates etc. given at the respective (national, regional or global) level, and, we assume, optimally from that particular point of view. A fundamental problem with this, in a climate policy context, is that such a view tends not to be correct, even when formulated at the national level. As a key feature of the GHG emissions control problem, a global view is needed, where the marginal externality cost at the *global* level is incorporated. Local decision makers are likely not to behave this way, except when international agreements dictate that globally optimal (emissions and energy) prices be applied. Little today indicates that such "optimal" prices will be applied in the near or intermediate future; thus a discrepancy between the ideal, global, social planner and the practical decision maker will be the order of the day. One objective of our analysis is to study how much such a decision maker is likely to deviate from a "socially optimal" decision (from a global optimality point of view).[3]

Increased availability of retrofit and closedown options affect expected fossil-fuel consumption and GHG emissions, in two opposing ways. For one thing, they reduce expected fossil-fuel consumption and emissions through the option to avoid such consumption and emissions ex post, in states where emissions and energy costs are particularly high, and instead retrofit or close the infrastructure down. On the other hand, an anticipated increased availability of such options serves to increase the chosen energy intensity embedded in the infrastructure. A greater availability of the two additional options make it less risky for the decision maker to choose a high initial (fossil-fuel) energy efficiency.

We find in simulations, in Section 7 below, that a higher variance on retrofit and/or energy costs (for given unconditional expectations of these costs) generally reduces expected future costs, both in terms of energy use, and in terms of total (energy plus retrofit) costs. With more uncertainty about both energy and retrofit costs (for given unconditional expectations), the retrofit option is exercised in more cases, and to greater benefit (as a higher variance implies a larger set of states where the retrofit option is "very gainful".[4] This is due to the general result that greater variances are advantageous, as they lead to the availability of more low-cost alternatives, which are exploited under an optimal ex-post policy.

The distribution of retrofit costs in period 2 depends on technological retrofit possibilities, which in turn are affected by any R&D effort to develop such technologies. Greater R&D effort, leading to reduced retrofit costs in period 2, increases the energy intensity of the initial infrastructure investment choice. Overall expected energy consumption may then either increase or decrease, depending on which of two effects is stronger: the increase due to greater energy consumption in "business-as-usual" states (resulting from the higher energy intensity); or the reduction due to the "business-as-usual" solution being chosen in fewer states (in period 2 of our model). Simulations indicate that the latter factor may in some cases dominate.[5] As a consequence in such cases, overall expected energy (and climate) costs are reduced with more energy technology R&D. Expected retrofit costs may also increase or

[3] See Strand (2009) for further elaboration of these issues.
[4] This conclusion holds when the decision maker is risk neutral; it may need qualification under risk aversion.
[5] Note the restriction on these conclusions, that the simulations to which we refer, are throughout based on the assumption that energy cost and retrofit costs in period 2 are lognormal, and independent.

decrease. The factor contributing to a decrease is the very drop in cost. The factor that could lead to an increase is that the retrofit option is exercised in more states of the world (thus avoiding energy expenditures entirely in such states). Typically here also, the first factor seems to dominate. Similarly, when the retrofit and closedown options are both available, an upward shift in the distribution of energy costs may increase or reduce expected energy costs, as well as expected retrofit costs; but, most likely (as apparent from our simulations), both will increase. Total expected costs then also clearly increase.

The analytical and quantitative literature dealing with such issues is small. Arthur (1983), David (1992) and Leibowitz and Margulis (1995) provide background by defining and discussing the issue of path dependency and its implications for future actions. The more specific topic of infrastructure choice and its implications for mitigation policy is discussed only recently. Shalizi and Lecocq (2009) provide a discussion of infrastructure costs and constraints which is more applied and intuitive than that provided here. The persistent effects of infrastructure choice on energy consumption and carbon emissions are discussed also by Brueckner (2000), Gusdorf and Hallegatte (2007a,b), and Glaeser and Kahn (2008). In particular, Gusdorf and Hallegatte (2007a) study the energy intensity of urban infrastructure for given population density. They focus in particular on inertia resulting from established urban structure, in response to “low” initial energy prices, which may later rise. They show, through simulations, that a permanent energy price shock leads to a transition period that is long (20 years or more) and painful (with high energy costs, and carbon emissions), but that energy consumption eventually will fall toward a substantially lower steady-state level. Glaeser and Kahn (2008) by contrast focus on energy consumption implications of differences in population density, both within urban regions and when comparing urban and rural population patterns. In this context they seek to quantify relationships between energy consumption and spatial patterns of cities in the U.S. They find, in particular, much lower per-capita energy consumption, and carbon emissions, in central cities than in suburbs. This bears on our analysis as it indicates that “compact” infrastructure (as found e g in central cities) is less energy demanding than “less compact” (found e g in suburbs).

The option to retrofit already established infrastructure, by removing either the initial energy requirement, or the carbon emissions associated with it (via CCS technology or replacing fossil fuels by renewables), may potentially reduce the inertia associated with the infrastructure. This is a focus in this paper, and also two further papers, Strand (2009) and Framstad and Strand (2009), which deal with complementary issues. Strand (2009) considers different utility function representations and their implications, in a similar setting. Framstad and Strand (2009) study optimal infrastructure investment when future energy prices follow a continuous stochastic process, where a delayed retrofit decision has a positive option value. Implications of retrofit possibilities and costs are further discussed analytically by Jaccard (1997) and Jaccard and Rivers (2007). The latter paper studies three types of demand-side infrastructure: urban structure; buildings; and equipment. The authors argue, based on simulations (and using a discount rate of 3 percent), that for buildings, and even more for urban structure, it is generally advantageous to make strong considerations for future emissions even when emissions prices start low and increase strongly over time (while this is often not the case for equipment where natural turnover provides sufficient flexibility). Shalizi and Lecocq (2009) provide a broader and more practically oriented discussion, with examples from both energy demand and supply; their overall argument is that energy-intensive infrastructure involving supply is generally more rigid than that involving demand; but sometimes (but not always) more prone to complete retrofit.

A further issue in the literature, related to the main themes of this paper, is the concept of a "low-carbon society" and ways to achieve it; see Strachan et al (2008a, b), and Hourcade and Cerassous (2008).[6] The overriding idea here is rather similar to that in the other literature cited: namely, that achieving a society with low GHG emissions (necessary for efficiency in the long run) requires a high concern for the design of current infrastructure investment. We also note that two World Development Reports, the WDR 2003 (World Bank (2003)), and the most recent WDR 2010 (World Bank (2009)), both have "inertia in physical capital" (echoing our analysis of infrastructure) as main themes.

2. Two-Period Model with Uncertain Retrofit Costs

Consider a world existing for two "periods". Infrastructure investment is made at the start of period 1, and can be "retrofitted" at the start of period 2.[7] As long as it is operated and not retrofitted, a given infrastructure gives rise to a given energy consumption per unit of time, determined at the time of initial investment. Energy supply costs and environmental/climate-related costs are uncertain at the time of establishment in period 1, but are revealed at the start of period 2. We assume that when retrofitted, the infrastructure is purged of all fossil-fuel energy content and/or all its carbon emissions. The infrastructure however still provides the same utility services to the public as it did before the retrofit. "Retrofitting", we assume, is not available in period 1: it represents a new technology, developed and available at the start of period 2.

Period 1 has unit length, while T is the "length" of period 2. T in principle may be given two alternative interpretations. First, it could simply be interpreted as the time elapsing during period 2, relative to the initial (unit) period. To invoke this interpretation in the following, it would need to be coupled with an assumption that decision makers choose a zero discount rate. Alternatively, T could embed discounting, in which case it would represent the discounted value of period 2 relative to that of period 1.[8] Under this interpretation, heavier discounting would lead to reduced T for given period length.

We also assume that the infrastructure can in principle be shut down at the start of period 2. Such action will be taken when the total utility of operating the infrastructure is less than the minimum of the energy cost of operation, and the retrofit cost, in period 2.

In period 1, the unit energy cost is q_1 (given and constant).[9] The policy maker decides on an infrastructure investment with given capital cost K. For simplicity and to focus on other issues than investment size, assume that all relevant infrastructure projects have the same investment

[6] A very early champion of this line of thinking and discussion was Amory Lovins; see, in particular, Lovins (1977).
[7] In the model as it otherwise stands, the assumption that a retrofit can be done only at the start of period 2, and not during this period, is no limitation as, we assume, no new information (nor any new or better retrofit technology) will be forthcoming during period 2.
[8] More precisely, when unity represents the present discounted value of a current income flow of one dollar throughout period 1, T would in this case represent the present discounted value of a current income flow of one dollar throughout period 2, as evaluated from the start of period 1.
[9] In the continuation, when we say "energy cost", we mean the combined energy and environmental cost associated with (fossil-fuel) energy use. This would be unproblematic when all environmental costs are charged to energy use in the form of energy taxes and quota prices. It is more problematic when this is not the case; this issue is elaborated more in the final section.

cost. Infrastructure type is in the model identified by a given energy intensity *H*, where we assume that all energy consumption associated with the infrastructure is fixed once the infrastructure is established, and until it is possibly retrofitted. Considering only economically viable projects, we focus on one particular trade-off only: an infrastructure project with higher energy content must give higher immediate utility, but will be more costly to operate over its lifetime due to its greater fossil-fuel energy requirement. Call the current (per time unit) utility flowing from the infrastructure when being operated $U(H)$, where $U'(H) > 0$, $U''(H) < 0$. We assume that $U(H)$ is given and constant and the same in both periods (and thus not subject to uncertainty).

Three alternative actions may be chosen in period 2:

1) *No new action, proceed with "business as usual*". In this case the full energy cost will be incurred in period 2. This is the optimal strategy when the energy cost in period 2 turns out to be lower than either the retrofit cost, or the average period-2 utility of the infrastructure.
2) *Retrofitting the infrastructure*. This is the optimal strategy when the retrofit cost in period 2 turns out to be lower than either the energy cost, or the average period-2 utility of the infrastructure.
3) *Infrastructure closedown*. This is optimal when environmental and retrofit costs both turn out to be higher than the average period-2 utility of the infrastructure. Note here however that closedown is a drastic measure. It will typically require that other infrastructure is supplied, to replace the services lost by project closedown. This is not explicitly modelled here. Implicitly, however, we may take our model to embed such effects, via the absolute value of the utility flow provided by the infrastructure (which should be defined relative to a situation where the utility is missing; thus a "relatively drastic" alternative).

The problem of a decision maker in establishing the infrastructure in period 1 is to select an energy investment intensity *H* so as to maximize

$$EW(1) = U(H) - q_1 H + EW(2) \tag{1}$$

where *E is* the expectations operator, and *W(2)* is the (optimized) value function associated with the infrastructure in period 2 (embedding the optimal action, among alternatives 1-3 above).

EW(2) embeds the decision maker's optimal responses at the start of period 2 (assuming that no further changes occur during period 2). Define $F(q)$ as the (continuous ex ante, when viewed from period 1) distribution over q levels to be realized in period 2, with support $[0, q_M]$, where q_M could be large.[10] Possibly, the *F* distribution for period 2 is shifted up by increased emissions in period 1.

We assume a perfectly continuous distribution over retrofit costs, y, in period 2 given by $G(y)$, with support $(0, y_M)$, where y_M could also be large.[11] Retrofit costs cannot be negative, but

[10] In simulations below we assume that F is log-normal, in which case F is not bounded above (it is however "thin-tailed").
[11] Thus total retrofit costs in period 2 are given as *Ty*.

could in principle be small in period 2, depending on the technology available for substituting out the fossil-fuel energy consumption or purging carbon from fossil fuels at that time. We assume that an infrastructure project, after a retrofit, incurs no energy costs nor other current costs in period 2, apart from the retrofit cost itself (which in the model is "periodized" in the same way as energy cost).[12] In the analytical presentation, period 2 realizations of energy cost and retrofit cost are assumed to be independent.[13]

Consider the choice between the three alternatives lines of action 1) – 3) in period 2. We start with action 3), project closedown. Define total utility per energy unit for installed infrastructure by $U(H_1)/H_1 = y^*$, where H_1 is energy intensity associated with the infrastructure investment chosen in period 1. Action 3) will then be chosen when the cost per unit of energy *q*, and the retrofit cost *y*, both exceed y^*. The probability of this event, when viewed from period 1, is

$$(2) \qquad P(3) = [1 - F(y^*)][1 - G(y^*)] .$$

Assume $0 < y^* < \min\{q_M, y_M\}$, and $0 < F(y^*), G(y^*) < 1$, implying $P(3) > 0$.

The probability P(1) of inaction (action 1), is given by the following expression:

$$(3) \qquad P(1) = \int_{q=0}^{y^*} [1 - G(q)] f(q) dq .$$

The probability *P(2)* of retrofit (action 2) is complementary (equal to 1 – *P(1) – P(3)),* but can also be found in a similar way as *P(1),* as follows:

$$(4) \qquad P(2) = \int_{y=0}^{y^*} [1 - F(y)] g(y) dy$$

The expected "per time unit" period 2 energy and retrofit costs as viewed from period 1, given an optimal strategy for period 2, are, respectively

$$(5) \qquad E[CH(2)] = \left(\int_{q=0}^{y^*} [1 - G(q)] f(q) q dq \right) H_1$$

[12] Alternatively, the retrofit cost could be interpreted to include some energy cost. This is unproblematic as long as the retrofit cost can be periodized.
[13] Independence of costs in this context is not obvious. It may however be considered as realistic in cases where the processes by which the two are affected, are quite different. But cases where the two costs will be correlated are also possible, and are often more realistic. Such correlations could be either positive or negative. Negative cost correlation may occur when the energy cost in period 2 is anticipated during the period of R&D efforts to develop new retrofit technologies. A high anticipated energy cost may then make the development of retrofit technologies more urgent, and more effort expended for this purpose. The two costs might then be negatively correlated. On the other hand, common drivers may affect both costs. This is relevant e g when energy cost is correlated with general production cost; when a retrofit involves some use of fossil energy, or the subsequent use of renewable energy whose marginal production cost is positively correlated with the cost of fossil fuels. In such cases the two cost variables would tend to be positively correlated.

(6) $$E[CR(2)] = \left(\int_{y=0}^{y^*} [1 - F(y)] g(y) y dy \right) H_1$$

E[CH(2)] expresses energy costs per "time unit" in period 2, while *E[CR(2)]* similarly expresses retrofit costs when similarly periodized (counted per period unit). H_1 denotes (per-period) energy consumption associated with the infrastructure as established in period 1.

Two alternative interpretations of retrofit costs have slightly different implications for the analysis. The first is to view retrofitting simply as replacing a fossil fuel with a non-fossil fuel which gives rise to no GHG emissions. Retrofit costs are then incurred currently in the same way as regular (fossil) fuel costs. Under the second interpretation, retrofit costs represent an investment to remove the fuel need tied to the infrastructure, or the emissions associated with the fossil fuel. In (6), this implies that *E[CR(2)]* must be "periodized" (given that T is different from unity), and spread evenly across T time units in period 2.

Expected (discounted) net utility from the infrastructure when being operated in period 2 is denoted EW(2), and equals the gross utility of the infrastructure, $TU(H_1) = Ty^*H_1$, in the states where it is not closed down in period 2 (thus with probability 1-*P(3))*, minus total combined expected energy and retrofit costs, *T{EC(2)} = T{E[CH(2)] + E[CR(2)]}*, over states where the infrastructure is operated (without, or with, retrofitting). We then have

(7) $$EW(2) = \{ y^*[1 - P(3)]H_1 - E[CH(2)] - E[CR(2)] \} T \; .$$

The first-period decision problem is formulated as maximizing the expected utility of the infrastructure investment in period 1, considering an optimal strategy in period 2. Define

(8) $$EW(1) = U(H_1) - q_1 H_1 + EW(2) = (y^* - q_1) H_1 + EW(2) \, .$$

Assume for now that the distribution of period 2 energy costs is exogenous (and not affected by emissions arising from the infrastructure). The solution to the maximization problem in period 1 takes the form

(9) $$\frac{dEW(1)}{dH_1} = U'(H_1) - q_1 + \frac{EW(2)}{H_1} + [U'(H_1) - y^*] \frac{d\left(\frac{EW(2)}{H_1} \right)}{dy^*} \, .$$

From (7) we find

(10) $$\frac{d \frac{EW(2)}{H_1}}{dy^*} = [1 - P(3)]T$$

Using the definition of *EW(2),* and setting the derivative in (9) equal to zero, we find the following implicit expression for the optimal energy intensity of the infrastructure:

$$U'(H_1) = \frac{q_1 + \dfrac{E(CH(2)) + E(CR(2))}{H_1}T}{1+[1-P(3)]T}$$

(11)
$$= \frac{q_1 + \left\{ \int\limits_{q=0}^{y*} [1-G(q)]f(q)qdq + \int\limits_{y=0}^{y*} [1-F(y)]g(y)ydy \right\} T}{1 + \left\{ \int\limits_{q=0}^{y*} [1-G(q)]f(q)dq + \int\limits_{y=0}^{y*} [1-F(y)]g(y)dy \right\} T} \cdot$$

$1+[1-P(3)]T$ equals the expected number of time units that the infrastructure will be operative during period 2. $\{E(CH(2))+E(CR(2))\}T = \{EC(2)\}T$ is the ex ante expected (energy plus retrofit) cost in period 2, which divided by H_1 is measured relative to energy intensity as established in period 1. $\{EC(2)\}T$ represents total energy costs plus retrofit costs during the period 2 expected operation time $[1 – P(3)]T$. The expression in the curled bracket in the numerator of (11) then denotes expected energy plus retrofit cost per unit of energy consumption defined by the established infrastructure. (11) can then be given a simple interpretation: the marginal utility of increased energy intensity associated with the infrastructure investment (the left-hand side of (11)) should equal the average energy plus retrofit costs incurred over the lifetime of the infrastructure capital (the right-hand side).

The optimal energy intensity is chosen according to average energy cost in operating the infrastructure, over its expected period of operation. Perhaps surprisingly, the extent of the operation period as such is not very important for the chosen intensity.[14] In e g a hypothetical case where the infrastructure is shut down in period 2 for certainty; the energy intensity would be determined simply by $U'(H_1) = q_1$. When average energy cost in period 2 is higher, $U'(H_1)$ is proportionately higher, and H_1 lower.

For y^* $(= U(H_1)/H_1)$ we find

(12)
$$\frac{dy^*}{dH_1} \equiv y_H^* = -\frac{1}{H_1}(y^* - U')$$

Here y^*-U' is positive and more so the more curved the utility function is in the neighborhood of H_1. Thus we can expect $y_H^* < 0$. This implies that a more energy-demanding infrastructure will have a lower threshold for operation, and will be "closed down" (or rather, replaced) ex post in more cases in period 2, when energy and retrofit costs increase. This may appear reasonable; remember that in our model all infrastructure projects are considered to be "equally large" in the sense of requiring the same initial investment cost. What distinguishes projects is the ex post energy requirement for their operation (or put otherwise, their energy intensity).

To consider a more specific example, take the simple case where the (vNM) utility function is quadratic:[15]

[14] This result is closely related to our initial assumption, that the size of the infrastructure investment is exogenously given.
[15] As usual the quadratic representation is a sufficiently precise (Taylor) approximation to any true utility function U as long as the changes in H_1 are not large.

(13) $$U(H_1) = aH_1 - bH_1^2$$

so that

(14) $$U'(H_1) = a - 2bH_1$$

Inserting into (12), this yields

(12a) $$y_H^* = -bH_1.$$

In this case, the absolute value of y_H^* is directly proportional to the curvature coefficient b in the utility function. Thus a relatively "flat" utility function implies a small change in the cut-off value y^*, for given H_1. A relatively "curved" function implies a large change in y^*.

Note also that with this specification

(15) $$U''(H_1) = -2b.$$

From (11), and given quadratic utility, the coefficient b plays a major role in defining the response of both H_1 and y^* to exogenous parameter changes. When b is small, H_1 will respond strongly and y* weakly to such changes. The opposite holds when b is larger. The interpretation of small or large b is thus of some importance. Intuitively, when b is small, the economy has several ways in which to design its infrastructure that are considered as "almost" equivalent for a given set of relevant parameters. Small parameter changes can then provoke a relatively large response for the "optimal" infrastructure as initially perceived.

Our concept of "optimality" is here defined with respect to the decision problem as set up. It is useful to identify different sources of inefficiency in deciding the initial infrastructure investment. We will distinguish between the following five points.

A) The initially expected distribution of energy costs, relevant for making current infrastructure decisions, is lower than (or more precisely, down-shifted relative to) the true (correct) distribution. This could occur e.g. when the infrastructure decision is base on an expectation of something close to current energy cost on average in period 2, while the correct distribution implies higher average energy costs.[16] In this case we may expect the infrastructure energy intensity to be chosen at a too high level, and fossil-fuel consumption in period 1 to be excessive. The period 2 realized expected energy consumption is however not obviously higher than optimal, as the closedown and/or retrofit options will be exercised in more cases (as we will see below).

B) The distribution of energy costs facing the policy maker is correctly anticipated, but is down-shifted relative to the "optimal" energy cost distribution. This case is relevant whenever the authorities, in the economy in question, implement emissions prices that are lower than "globally correct" prices. We will argue that it is a highly relevant case:

[16] We argue that this could occur even in cases where the entity making the infrastructure decisions would face a "correct" overall energy price (including the "true" costs of emissions). One such case is where the administrative procedure for making public investments involves incorporation of future costs and benefits for only a limited period (say, 20 years), while an appropriate investment decision would need to involve a much longer period (say, up to or exceeding 50 years). An increasing future energy price (also beyond the 20 year term) would add to the bias involved in such an investment procedure.

as of today, hardly any country implements what most analysis would agree are "globally correct" emissions prices; nor seems to be willing to do so in the foreseeable future. In this case we can expect, unambiguously, excessive fossil-fuel consumption in both periods.

C) Future retrofit costs are incorrectly anticipated. Such costs are likely to depend on the rate of technological progress for retrofit technologies, and are inherently unknown at the investment stage in period 1. A too optimistic view of this development so that the anticipated retrofit cost distribution is lower than the correct distribution will then lead to bias in the direction of excessive energy intensity for the initially established infrastructure. Note however that when the distribution of future energy costs is too optimistic, the retrofit cost distribution will be less important for the energy intensity decision, since the prior expectation is then that retrofits are necessary in fewer cases. On the other hand, instead of "technology optimism" one could have "technology pessimism" (a too pessimistic view of the distribution of future retrofit costs), which would work in the opposite direction.

D) Future retrofit costs are correctly anticipated, but are higher than socially optimal costs. This could be the case when there is a too low R&D effort in developing new energy technologies, including those for retrofits. This would leads to a higher than optimal distribution of retrofit costs. This would increase overall implementation costs (in particular, the minimum of energy and retrofit costs) in period 2. Given correct perceptions in period 1, the choice of energy intensity of the infrastructure is now *lower* than optimal (efficient energy intensity is greater than the one chosen), while the probability of ex post business-as-usual operation of the infrastructure is *higher* than optimal (as there are more states where the energy cost is below retrofit cost). On balance the second factor is likely to dominate, and, thus, socially excessive energy consumption in period 2. In such cases, and assuming no other distortions, the energy intensity of the initially chosen infrastructure would, in fact, be *suboptimal*. On the other hand, ex post energy costs in period 2 for given infrastructure are *overoptimal*. This follows straightforward from the fact that the business-as-usual alternative will be chosen in more cases in period 2, and the retrofit alternative in fewer cases.

E) The policy-relevant value of T, call it T_1, is less than the optimal value, call it T_0. Reasons for this may be either excessive discounting (for the case where T is interpreted as a discounted value), or that the initial policy decision undervalues the length of period 2. Since T, as noted at the start of this section, can be interpreted as a discounted value of period 2 relative to period 1, we may have a discrepancy between the "socially correct" value T_0 and the value T_1 used when deciding on H_1. $T_1 < T_0$ could then reflect excessive discounting. When average costs per operation period are greater in period 2 than in period 1 (as might be expected), this leads to a lower average operation cost for the infrastructure, and a more energy-intensive infrastructure.

In four of these cases (all except D), the overall expected fossil-fuel energy consumption (and GHG emissions) over the potential lifetime of the infrastructure is excessive from a social point of view, in the sense that it is higher than the expected fossil-fuel consumption and emissions for the case where all global externalities are optimally considered and anticipated.

3. Impacts of Shifts in the Energy Cost Distribution

We will now study some implications of changes in the distribution of energy costs in period 2. Consider then a downward shift α in the distribution of energy costs, that leaves all other parameters unchanged. Call the new distribution $F_1(q) = F(q+\alpha)$ (so that F_1 is shifted up relative to F by a constant amount α for any given q). This is formally the same as the entire distribution of q being shifted down, but retaining the distribution function F. Energy costs, in consequence, fall on average. In particular, since the distribution G is unaltered, the following new definition of *E[CH(2)]* then applies:

(5a) $$E[CH(2)] = \left(\int_{q=0}^{y^*} [1 - G(q)] f(q+\alpha) q dq \right) H_1$$

(6a) $$E[CR(2)] = \left(\int_{y=0}^{y^*} [1 - F(y+\alpha)] g(y) y dy \right) H_1$$

Differentiating the expressions for *P(1), P(2), E[CH(2)]* and *E[CR(2)]* with respect to α then yields (assuming that y^* is not significantly altered):[17]

(16) $$\frac{dP(1)}{d\alpha} = \int_{q=0}^{y^*} [1 - G(q)] f'(q) dq \, .$$

(17) $$\frac{dP(2)}{d\alpha} = -\int_{y=0}^{y^*} f(y) g(y) dy$$

(18) $$\frac{dE[CH(2)]}{d\alpha} = \left(\int_{q=0}^{y^*} (1 - G(q)) f'(q) q dq \right) H_1 + \frac{E[CH(2)]}{H_1} \frac{dH_1}{d\alpha}$$

(19) $$\frac{dE[CR(2)]}{d\alpha} = -\left(\int_{q=0}^{y^*} f(y) g(y) dy \right) H_1 + \frac{E[CR(2)]}{H_1} \frac{dH_1}{d\alpha} \, .$$

$dH_1/d\alpha$ is found totally differentiating (11) with respect to H_1 and α.

P(1) here increases (as *f'* is positive for low *G* values): the probability of "business as usual" increases. This is intuitive: when energy costs fall, the likelihood that the (business-as-usual) energy cost option is exercised in period 2 increases, "everything else equal". The probability that the retrofit option is exercised in period 2 drops unambiguously. The increase in the former is greater so that *P(1)+P(2)* increases (i.e. the closedown option, is exercised in fewer cases). Ignoring first effects via changes in H_1, we find that the effect of a shift in α on unit energy costs (represented by the first term on the right-hand side of (18)) is ambiguous. Two factors go in different directions: a greater *P(1)* implies that energy costs are incurred in more states, leading such costs to increase. On the other hand, unit energy costs drop for any given

[17] This requires that the coefficient *b* in (12a) is small, and that H_1 changes "relatively much" compared to y*.

state, which tends to reduce costs. The effect on expected unit retrofit costs is however unambiguously negative. This is intuitive: the only thing that happens to retrofit costs is that such costs are applied in fewer states, thus reducing overall expected retrofit costs.

A positive shift in α shifts overall unit energy plus retrofit costs down. From (11), H_1 then increases. Intuitively, lower overall expected operating (energy plus retrofit) costs in period 2 leads to selection of infrastructure with a higher energy intensity. The response of H_1 to changes in unit costs could in principle be large. Overall expected energy and retrofit costs could then easily increase when the distribution of energy costs shifts down, and this is correctly anticipated in period 1. Quantitative effects here require specifying functional forms in more detail.

For the particular case of a quadratic utility function (12a) in infrastructure energy intensity, y^* falls when H_1 increases. (12a) and (18) however show that when the response of H to cost is high (and b is low), y^* changes little in response to energy cost changes.

4. Impacts of Shifts in the Retrofit Cost Distribution

This section considers some impacts of changes in costs of retrofitting in period 2. By this we mean to study impacts on outcomes, from both marginal changes in retrofit costs, and from the retrofit option being at all available.

(5b) $$E[CH(2)] = \left(\int_{q=0}^{y^*} [1 - G(q+\beta)] f(q) q dq \right) H_1$$

(6b) $$E[CR(2)] = \left(\int_{y=0}^{y^*} [1 - F(y)] g(y+\beta) y dy \right) H_1$$

Differentiating the expressions for *P(1), P(2), E[CH(2)]* and *E[CR(2)]* with respect to β then yields

(20) $$\frac{dP(1)}{d\beta} = -\int_{y=0}^{y^*} f(y) g(y) dy$$

(21) $$\frac{dP(2)}{d\beta} = \int_{q=0}^{y^*} [1 - F(q)] g'(q) dq \,.$$

(22) $$\frac{dE[CH(2)]}{d\beta} = -\left(\int_{q=0}^{y^*} f(y) g(y) dy \right) H_1 + \frac{E[CH(2)]}{H_1} \frac{dH_1}{d\beta}$$

(23) $$\frac{dE[CR(2)]}{d\beta} = \left(\int_{q=0}^{y^*} (1 - F(q)) g'(q) q dq \right) H_1 + \frac{E[CR(2)]}{H_1} \frac{dH_1}{d\beta}$$

Interpretations are in this case similar to those in the case of energy cost changes. When β increases, the distribution of retrofit costs is (in analogous fashion to the energy cost distribution, in Section 3) is shifted downward, and average retrofit costs fall. The probability of "business-as-usual" energy consumption ($P(1)$) then decreases unambiguously, while the probability of retrofit ($P(2)$) increases unambiguously. The increase in the latter is also now in general greater, so that $P(1)+P(2)$ increases. Thus expected energy intensity of the infrastructure falls unambiguously (as represented by the integral on the right-hand side of (22)), while the change in expected retrofit cost per established energy unit is ambiguous (the integral on the right-hand side of (23)). Their sum, $EC(2)$, falls unambiguously.

Consider an alternative case where the retrofit option is no longer available (NR denoting the "no retrofit" case)[18]. We have the following probability of closedown in period 2:

(24) $$P_{NR}(3) = 1 - F(y^*), 1 - P_{NR}(3) = F(y^*) .$$

In this case the probability of (energy-demanding) infrastructure operation in period 2, $P_{NR}(1)$, is given simply by $1-P_{NR}(3)$. For given y^*, the probability of closedown is smaller when the retrofit option is available, $(P(3))$, than when it is not $(P_{NR}(3))$, by a factor $(b-y^*+y_0)/b$. The probability of operation (with or without retrofit) is correspondingly greater when a retrofit option is available. The probability of energy-demanding operation is smaller with the retrofit option, by a factor $(1-(y^*+q_0)/2b)$.

To study how a lack of retrofit option changes the initial energy intensity of the infrastructure, H_1, assume a further simplified case with no energy cost in period 1 $(q_1 = 0)$. The objective is then simply to compare the expected per-unit combined energy and retrofit cost in period 2 in the two cases. This cost equals $(EC(2)/H_1)/(1-P(3))$ in the case where the retrofit option is included, and $(EC_{NR}(2)/H_1)/P_{NR}(1)$ in the case where the retrofit option is not included. These are the respective expressions for average overall costs per operation time, or probability of operation in period 2. In either case this expression is to be set equal to $U'(H_1)$, for an optimal H_1 level to be achieved.

5. The Value of the Closedown Option

We now study the way in which the very availability of the closedown option affects the overall solution, initial energy intensity, and cost variables. The closedown option could be viewed as "unavailable" when the total utility per unit of energy consumed for the chosen infrastructure, $y^* = U(H_1)/H_1$, is so high that closedown is never a realistic option (i.e., $[1-F(y^*)][1-G(y^*)]$ is "very small"). It is here easiest to think of cases where the infrastructure involves a high sunk cost relative to energy consumption (such as, perhaps, for urban structures including housing and transport systems). We may then consider the limit as y^* tends to infinity.

Expected energy consumption in period 2 is then simply $P(1)$ multiplied by TH_1. Availability of the retrofit option, for given H_1, now reduces expected energy consumption by $P(2)TH_1$ in period 2.[19]

[18] One interpretation of such a case is that the lower bound of the retrofit distribution, y_0, is higher than the average total value (per unit of energy consumed) of the infrastructure in period 2.

[19] Of course, H_1 will not in general be given but determined endogenously; H_1 will tend to be greater when the retrofit option is available since expected costs are then lower.

The expected energy and retrofit cost, and ex ante probabilities of "business-as-usual" operation and retrofit in period 2 are now still given respectively by the general expressions (2)-(6). Note however that we now have *P(1) + P(2)* = 1. Expected ex ante utility of second-period operation is now

(7a) $$EW(2) = \{y^* H_1 - E[CH(2)] - E[CR(2)]\}T \; .$$

The first-period decision problem can now be formulated as maximizing *EW(1)*, from (8). The resulting solution for optimal energy intensity of the infrastructure is now found from the following condition:

(11a) $$U'(H_1) = \frac{q_1 + \frac{E[CH(2)] + E[CR(2)]}{H_1} T}{1+T} \; .$$

This can now be compared to costs when closedown is an option. There are two main differences between (11) and (11a). First, the respective expressions *E[CH(2)]* and *E[CR(2)]* are now greater, relative to the ex ante probability of operation and even more absolutely. Secondly, the term *[1-P(3)]* in the denominator of (11) has vanished. As a result, overall expected costs are greater, and the weight to second-period costs versus first-period costs is greater. Thus, when expected second-period costs "per period" are greater, this also tends to increase the overall expression on the right-hand side of (11). Overall, $U'(H_1)$ is increased, and H_1 reduced. Having an effective closedown option thus increases the energy intensity of the original infrastructure investment, relative to the case with no such option. This is of course unsurprising and intuitive: when the option to close down is available, it will be used only in states where both retrofit and energy costs are very high. Exercising the closedown option eliminates costs in these most expensive operation states, which in turn provides incentives to raise the infrastructure's initial energy intensity.

Notably, energy consumption and expenditure related to the infrastructure in period 2 are lower when the closedown option is available, than when it is not. This conclusion may however be deceptive: we have here assumed that when closing down no energy expenditure is incurred whatsoever. This is unrealistic since the closed down infrastructure will need to be replaced by an alternative that will likely demand energy (although presumably less than that initially established; in particular since replacements tend to occur in states with high energy costs in period 2).

6. Endogeneity of Retrofit Costs

The retrofit options available in period 2 are likely to follow at least in part from technology developments over period 1, and these may in turn be influenced by R&D efforts. The idea in this section is to study possible effects of such efforts.

Influencing R&D efforts with the purpose of mitigating GHGs has emerged as a core theme in the climate policy debate, from several angles. One is how an optimal climate and energy policy (in the form e g of emissions or energy taxes) can depend on the presence of R&D; this has been discussed e g by Goulder and Schneider (1999), Goulder and Mathai (2000), Bonanno et al (2003), and Greaker and Pade (2008). A separate issue is that while it may be very difficult to reach an international agreement to effectively reduce GHG emissions, using policy instruments such as emissions taxes and caps, some analysts claim that reaching an agreement to support emissions-reducing technological progress may be easier.[20]

Here we simply assume that the "decision unit" that carries out the initial infrastructure investment, may also carry out R&D activity to affect the options, and costs, of retrofitting this particular infrastructure in period 2. Assume that such investment only affects retrofit costs for this particular infrastructure, and not for other units nor more generally. We assume that the entire distribution function for period 2 retrofit costs can be given a constant vertical shift (in similar fashion as in section 4 below) through additional R&D effort in period 1.[21] This upward shift in distribution is the same as a downward cost shift, as in section 4, and was there given from (23), as a result of changes in a shift parameter β for this distribution.

Here, consider the following modified discounted utility as viewed from period 1:

$$EW(1;R) = (y^* - q_1)H_1 + EW(2) - R \tag{25}$$

where *EW(2)* is given from (7), *R* is the first-period R&D cost, the G function is shifted, with shift parameter β, and where the size of β is a positive function of *R*. We then derive the following general optimality condition with respect to *R*:

(26)

$$\frac{dEW(1;R)}{dR} = \left\{ \left[y^* - q_1 + y^*(P(1)+P(2))T \right] \frac{dH_1}{d\beta} + y^* T H_1 \frac{dP(1)+dP(2)}{d\beta} - T\frac{dEC(2)}{d\beta} \right\} \beta_R - 1 = 0$$

where β_R is the derivative of β with respect to *R*.[22] The partial derivatives in (26) are found from (20)-(23) plus (11) differentiated. While (26) looks complicated, its essence is that the total derivative of *EW(1)* with respect to *R* consists of three marginal benefit terms inside the curled bracket, classified by how model variables are affected: 1) effects via the increase in *H*; 2) via increase in the joint probability of period 2 operation, *P(1)+P(2);* and 3) via operation costs (energy costs plus retrofit costs) in period 2. These three terms are traded off against, and at the optimum set equal to, the unit cost of R&D investment.

An important parameter here is β_R. Presumably, the marginal effect of R&D costs diminishes with greater costs (as will, rather generally, be required for a unique internal optimum to be found under the problem (26)).[23] Consider here again a standard second-order Taylor expansion (and thus quadratic formulation)

[20] See in particular Barrett (2006, 2009).

[21] This is of course highly unrealistic. In practice, R&D efforts will affect retrofit costs more generally, and also for other projects, and thus imply positive externality effects for the latter. This issue is not discussed fully here. For further discussion see e g Golombek and Hoel (2005, 2006).

[22] A second-order condition here also needs to be fulfilled. A sufficient condition here is that β_R is decreasing in *R*.

[23] One way to visualize this is to consider R&D projects carried out in sequence by their likelihood of success; when only a few projects are funded these are the most promising.

(27) $$\beta(R) = \lambda R - \mu R^2$$

where λ and μ are positive constants, so that the first- and second-order derivatives of the β function are given by

(28) $$\beta'(R) = \lambda - 2\mu R > 0$$

(29) $$\beta''(R) = -2\mu < 0$$

The main point here is that when the curled bracket in (26) is large, the first derivative $\beta'(R)$ will be small, and R correspondingly large. A large overall positive utility effect of a given shift in the retrofit function then leads to a large optimal R&D effort R.

We have here assumed that the investment in question only affects costs for one particular infrastructure facility. In practical cases of individual R&D projects that affect future infrastructure costs, such R&D expenditures are likely to have most of their effects on costs for other projects.[24] In the context of our model, the marginal social benefit of R&D is much greater than the "private" benefit for the infrastructure project sponsor. Consider here a case where (27) correctly represents the overall social impact of R on retrofit costs, while the private impact is only a fraction h (< 1) of the social impact. In this case, the marginal change in β when R changes, as perceived privately, is also a fraction h of the social impact given by (28), and thus

(28a) $$\beta'(R;h) = h(\lambda - 2\mu R) > 0.$$

When h is smaller, R must be smaller to fulfil (26).[25] As a result, the R&D activity will be (perhaps much) lower than optimal when most of the overall returns to private R&D accrue to others. One then faces an obvious problem of policy coordination across countries, which in principle could be as serious as that for regular mitigation policy. High appropriability of rents to developers of new technology will tend to reduce this coordination problem.[26]

7. Simulations

We will now illustrate some basic properties of this model, through simulations of energy and retrofit cost distributions, under some simplified assumptions. The simulations will focus on the period 2 cost structure, and how conditional energy, retrofit and total expected cost is affected by expectation and variance of the two cost items, given an optimal ex post strategy to minimize these costs.[27] Most of these simulations depart directly from the expected cost expressions (5) (expected energy cost in period 2) and (6) expected retrofit cost in period 2). We generally take H_1 to be given (i e we do not study the optimal H_1 decision) and equal to

[24] Another way to express this effect is that there is likely to be a high degree of "technology spillovers" associated with R&D for development of new retrofit technology; see discussions of such spillovers e g by Golombek and Hoel (2005, 2006).

[25] With this formulation, when h is small, no such non-negative R can be found. Then no R&D investments will be undertaken by private agents.

[26] One way of securing this is strong patent laws. But this has other negative side effects, in particular, the markets in which the newly developed technologies are applied will not be competitive; see e g Greaker and Pade (2008).

[27] Simulations have been done using Matlab.

unity (in other words, the basic energy intensity of the infrastructure is unity), and thus study changes in unit energy and retrofit costs only.

In the simulations, if nothing is otherwise stated, the (unconditionally) expected energy and retrofit cost are both kept constant, setting $E(q(2))$ (= $\int gf(q)dq$) = 2, and $E(y(2))$ (= $\int yg(y)dy$) = 3. ($E(q(2))$ is the (unconditional) expected energy/environmental cost in period 2; it would be the *actual* expected cost given no retrofit or closedown. A similar interpretation holds for $E(y(2))$.) While both energy and retrofit costs are uncertain, energy costs are thus, in the "benchmark" case, assumed to be lower *in expectation*. The distributions of energy and retrofit costs, $F(q)$ and $G(y)$, are both assumed to be log-normal, and independent.[28] The cut-off level for costs (beyond which the infrastructure will be abandoned) is, unless otherwise is stated, set at y^* = 10 (= 5 times unconditionally expected energy costs), for all simulations.

Figure 1 describes conditional expected energy costs (blue), retrofit costs (green), and total costs (red), and how they vary with changes in the variances of energy cost (q) and retrofit cost (y). Thus, under certainty no retrofits would ever take place in period 2 (since retrofit cost would be higher than the costs of normal, non-retrofitted, operation), and period 2 cost would simply equal 2. Under uncertainty, additional options open up, as particularly high (energy, and retrofit) costs can be avoided, and cases with low costs implemented. As a result, overall conditional costs will be lower, and more so the greater is uncertainty, represented here by the variances of q and y. In the figure, this feature is seen to hold for partial increases in both variances. In particular, when both variances are about 2.5, approximately half of (unconditional) energy cost is avoided, while half as much is added in the form of retrofit cost. The total overall factor cost saving is then about one fourth. Note also, as a general feature of the results from the simulations, that the factor with the lower *unconditional* expected cost has the higher *conditional* expected cost. The reason is, obviously, that when the unconditional expectation is lower, the respective alternative will be applied in more cases (and the opposite alternative in fewer cases).

Equally important here is the ability to avoid energy costs as such, as a function of increased uncertainty. We see from Figure 1 that, in many cases, at least half of potential energy cost is avoided ("conditional *Eq*" being less than unity).

Figures 2 and 3 repeat similar exercises except for cases where the respective variances are, respectively, much lower, and much higher, than those shown in Figure 1. From Figure 2, when the variances are very low, the solution tends to one with only the basic energy cost option with leads to an energy cost equal to 2, and no retrofit cost (as there are very few cases where the retrofit alternative outcompetes the energy cost alternative). We see however, in the lower figure of Figure 2, that there is some more substantial substitution of retrofit cost for energy cost when the variance on y (retrofit cost) increases beyond 0.2. Overall cost however remains close to 2 in these cases.

[28] The independence assumption can be questioned, as noted above, and will be altered in subsequent work.

Figure 1: Expected energy, retrofit and total costs, and the variance of *q* and *y*.

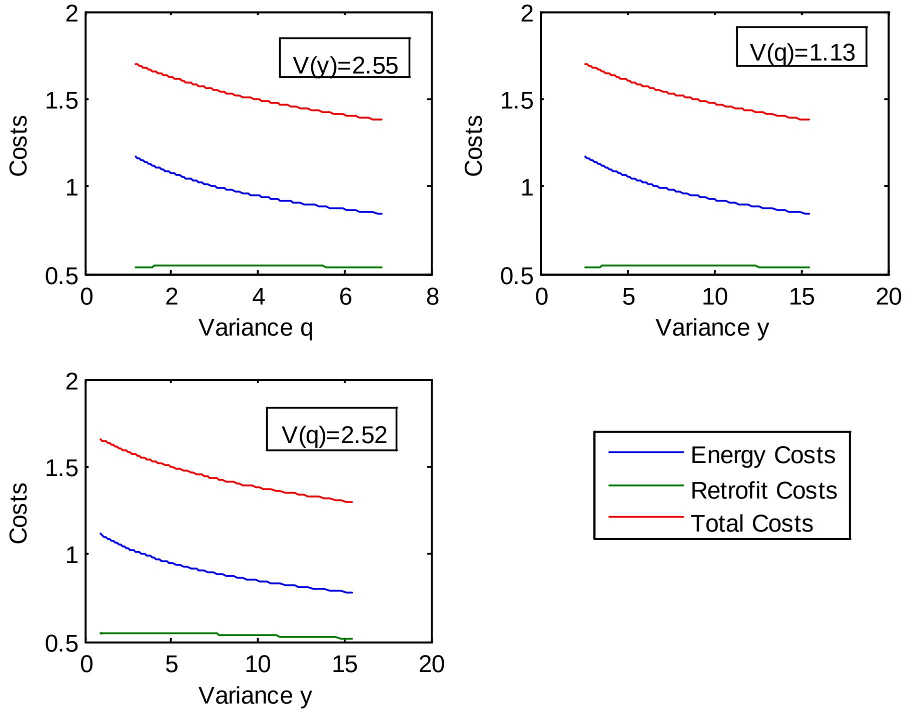


Figure 2: Expected energy, retrofit and total costs, and the variance of *q* and *y*. (Low variances)

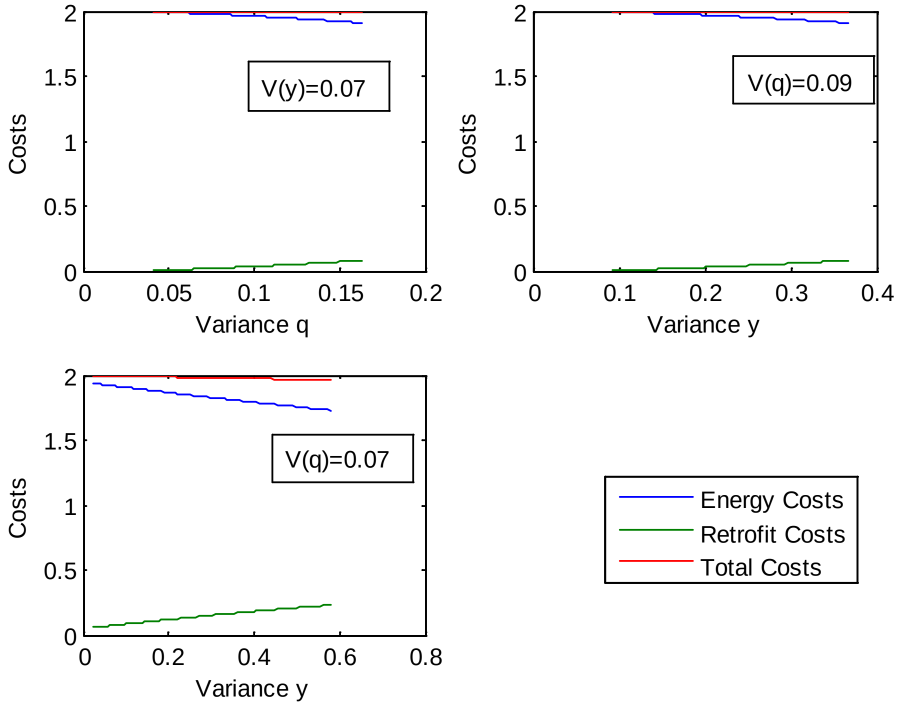
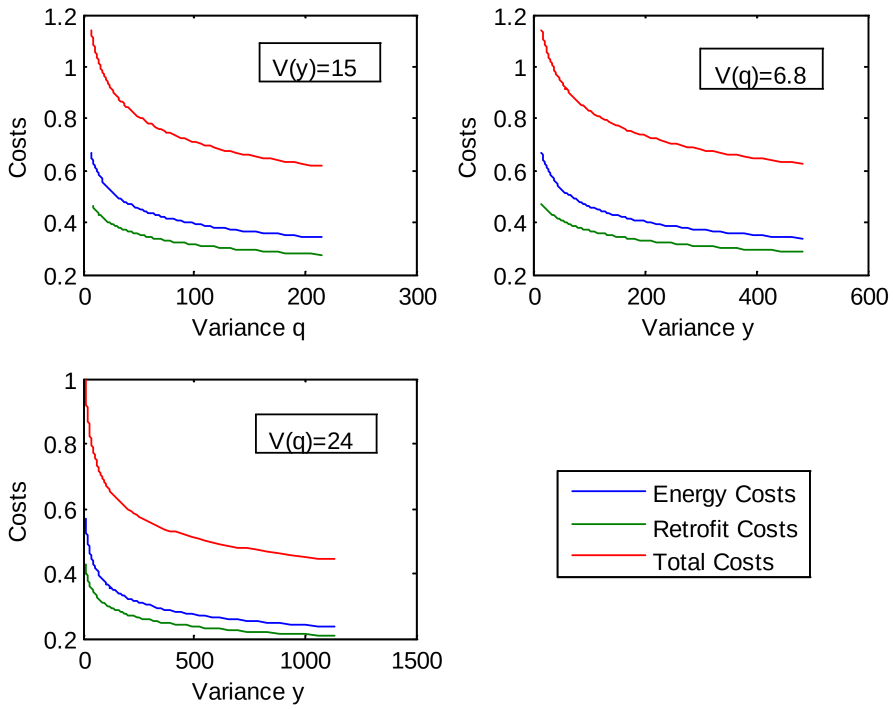


Figure 3: Expected energy, retrofit and total costs, and the variance of *q* and *y*. (High variances)



In Figure 3 the situation is rather different. Here variances on both *q* and *y* are assumed to be much larger than in the Figure 1 alternative. The scope for substitution of the two factors, and the corresponding scope for cost avoidance, is then much greater. Both energy and retrofit costs are here incurred about proportionately in the various alternatives, and more than half of potential total cost is avoided in many cases; and in some cases as much as 75 percent of total potential energy cost is avoided. Clearly, the amount of energy cost that is effectively tied up by the initial infrastructure investment is then rather small.

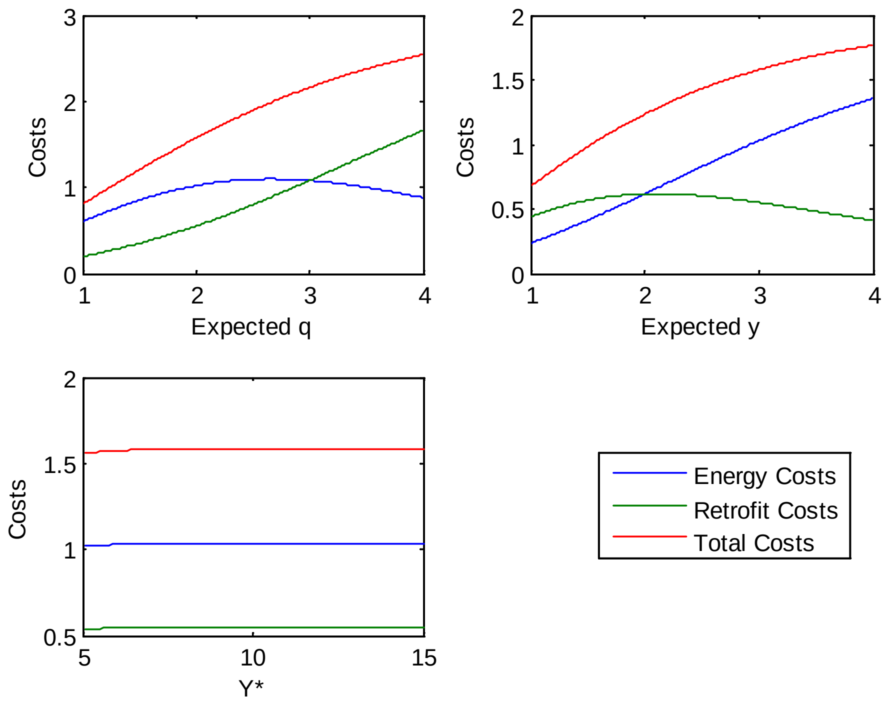
Figures 4-6 show an alternative set of simulations, where the expectations of energy and retrofit costs are changed parametrically, for given variances. In Figure 4, variances are in an "intermediate" range, both equal to 2.5 (and again, when nothing else is indicated, $Eq = 2$, and $Ey = 3$). We find, as already commented, that an increase in expectation for one cost item leads to an initial increase (roughly, up to the expected value of the other cost item) and subsequently a decrease in conditional expected cost associated with that item. We also see that, at most, about one third of the overall potential cost is avoided by optimal substitution given these variances (when both expectations are equal; either both are 2 or both are 3).

Figure 5 deals with the case where the variances are "low". Overall, there is then little scope for cost avoidance; total expected cost is very close to the overall lower cost alternative. There are here also more dramatic changes in conditional expected *q* and *y* when expectations increase beyond the value of the opposite cost parameter; this is of course due to the almost complete phasing out of retrofit cost whenever the expectation exceeds that for energy cost; and vice versa.

In Figure 6, by contrast, variances are "high". In this case, there is no apparent strong tendency for either one of the cost items to be phased out when its expectation increases. The reason is that, even when the expected energy (retrofit) cost is high, there is still a high probability that the actual energy (retrofit) cost is low, and at the same time the retrofit (energy) cost higher, leading to energy use (retrofit) being the chosen application.

The last (bottom) tables in each of Figures 4-6 show overall effects of changes in closedown costs; or rather effects of the possibility to exercise the closedown option. We assume that the cut-off value for costs (beyond with the closedown option will be exercised) is non-stochastic but is varied parametrically in the table. As can be seen from the figures, for "reasonable" values of the cut-off level *y** (beyond 5; a quite low level),[29] the closedown option has consequence for overall costs only in the high variance alternative (Figure 6); and seemingly not at all in the two other alternatives (this is seen by the "total cost" curves not being sensitive to *y**). The reason is, obviously, that in the (low- and moderate-variance) alternatives, the closedown option will be exercised in few cases; and, when exercised, on average lead to only moderate cost savings. This is different in the high-variance alternatives (Figure 6), where the "total cost" curves are obvious (increasing) functions of *y**. The simulations here indicate an overall (expected energy plus retrofit) cost saving of about 10 percent when the cut-off level *y** is lowered from 15 to 5) (and a similar relative cost saving for energy costs alone).

Figure 4: Expected energy, retrofit and total costs, and the expected values of *q*, *y* and *y.**



[29] Note that, for *y** = 5, the *overall net utility value* of the infrastructure *after* infrastructure costs have been sunk is 2.5 times *ex ante expected* energy costs; which seems quite moderate.

Figure 5: Expected energy, retrofit and total costs, and the expected values of *q*, *y* and *y. (Low variance of *q* and *y*; *V(q)*=0.07; *V(y)*=0.09)**

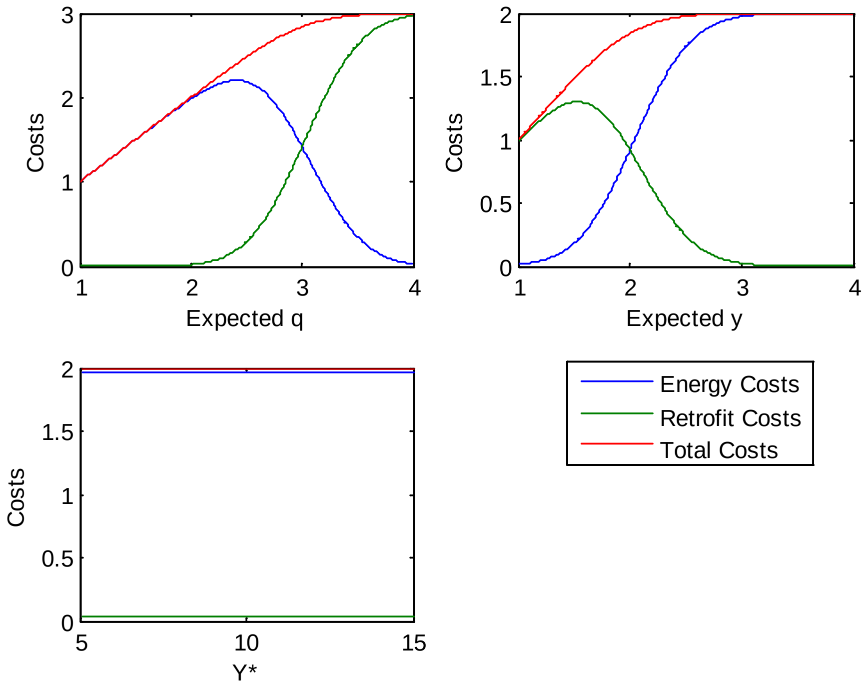
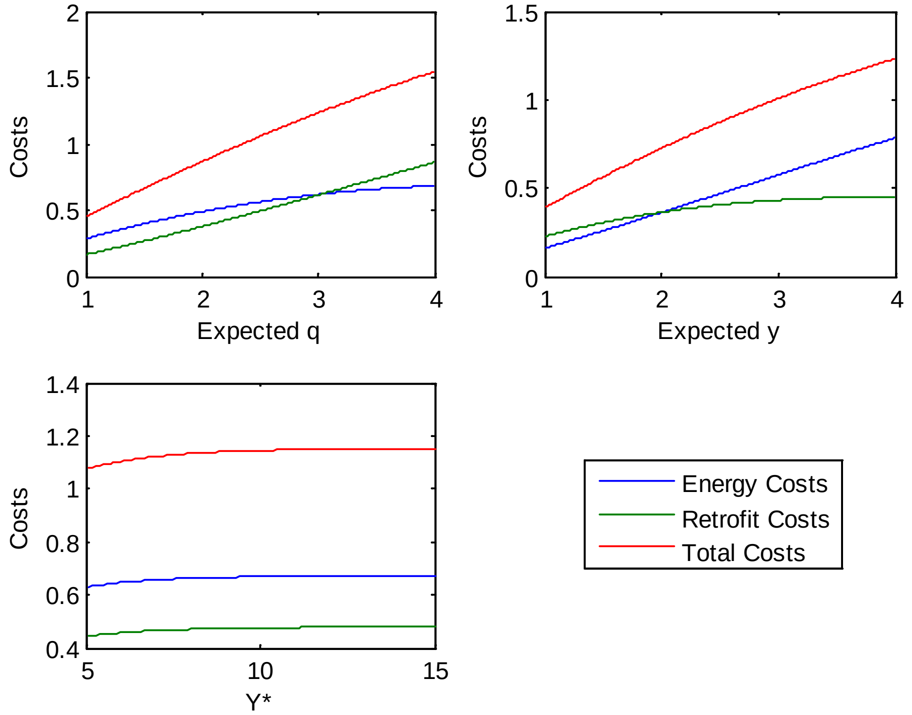


Figure 6: Expected energy, retrofit and total costs, and the expected values of *q*, *y* and *y (High variance of *q* and *y*; *V(q)*=7; *V(y)*=15)**



Some striking results follow from these simulations. While the expectation of *unconditional* energy cost *E(q(2))* equals 2, the *conditional or actual* expected energy cost related to the infrastructure in period 2 is lower in all cases. The difference is greater when variances (of both *q* and *y*) are larger. This is due to the decision maker optimally exercising either of our two other options ex post, retrofitting or closing down. Energy costs are avoided in states of the world where they are particularly high, and also in states when they are low but retrofit costs are even lower. These effects are stronger, the more variable energy and retrofit costs are (for given unconditional expectations).

Concentrating on energy costs alone, (conditional) expected energy costs can be reduced when costs become more variable (for given unconditional expectations), for three separate reasons. It is then more likely that a) a given retrofit cost is lower, and that b) a given utility value of continued operation is lower than actual realized energy cost. Besides, c) a more variable retrofit cost increases the likelihood that the retrofit cost (with given expectation) is lower than any given energy cost. All these factors tend to ameliorate the overall effect of the initial "tying up" of energy costs associated with a given infrastructure.

The figures illustrate a further feature of the theoretical analysis, namely that when expected (*unconditional*) retrofit cost is greater than expected energy cost (for given *var(q)*), expected conditional or *actual* energy cost is greater than expected retrofit costs. This is because when a given expected cost is high, the respective alternative tends to be exercised in fewer cases.

A few cautionary notes must be recognized when interpreting these simulations. In particular, the model has the (perhaps unrealistic) feature that when the closedown option is exercised, there is neither energy consumption nor any emissions. The same holds when the retrofit option is exercised. This assumption could however, without much loss of generality, be weakened by assuming a certain (minimum) level of energy use and emissions in this case.

Another reservation is that the distributions of energy and retrofit costs are assumed to be independent.[30] Perhaps more likely, these costs are in practice positively correlated (when one is high, it is more likely than otherwise that the other is high, etc.). Positive correlations lead, in general, to less scope for cost savings.

Note also that these simulations do not directly address a key issue for the overall analysis, namely, the effect on energy intensity of the initial infrastructure investment. The general conclusion, from section 2 above, is that having options (through retrofit and closedown possibilities) to reduce energy use later increases the energy intensity of the infrastructure. We have no general conclusions on the strength of this effect, but it could be substantial, in particular when infrastructure investment choice (unlike in our model) implies a choice between two discretely different infrastructure systems (such as between a transport system "largely" dependent on private, or public transport), and where the policy maker is initially "almost indifferent" between the two.

8. Summary of Sources of Market Inefficiency, and Final Comments

A main aim of this paper has been to study decisions to invest in infrastructure that commits society to potentially high levels of energy use, and carbon emissions, for a long future

[30] The log-normal distribution assumption could, potentially, also be attacked. Log-normality is however a rather robust assumption in this context; see e g Schuster (1984).

period. We have studied factors behind inefficiency of such investments, and implications for GHG emissions (which would be excessive in the long run), and considered how any inefficiencies can be avoided or counteracted. Another aim of the paper has been to study the impacts of two types of policy interventions that may be applied after infrastructure investment have been sunk: namely first, "retrofitting" the infrastructure (by making an additional, later, investment that removes the energy demand and/or emissions due to the infrastructure, while retaining its utility value to the public); and secondly, closing it down (a more drastic alternative, as the utility value of the infrastructure is then is removed together with energy use and emissions). Most of the focus here has been on the retrofit alternative, on its ability to reduce subsequent (energy and environmental) costs, and its effect back on energy intensity of the initial infrastructure.

Considering the first of these objectives, inefficient infrastructure choice can result from all the types of market failure discussed in section 2 above. The five types discussed were A) anticipated energy and emissions costs below actual prices; B) too low actual energy (including emissions) costs facing private agents, set by the respective governments; C) incorrectly anticipated (and lower than actually realized) retrofit costs; D) too high realized retrofit costs; and E) excessive discounting.

Roughly, these explanations can be classified into two groups: one related to insufficient or faulty general climate-related or energy policies (including insufficient emissions pricing and technology support); and another related to inefficiencies and incompleteness in the execution of policy. Points B and D fall largely into the former category, while points A and C mainly into the latter. Point E might conceivably fall into either category.

One easily understands the main reason for faulty or inadequate policies of the individual governments hosting infrastructure projects: namely, the basic lack of incentives of governments to address the problems of mitigation, at least in the absence of a comprehensive and binding agreement for so doing. The second group of explanations has a more diverse set of explanations, which are however all related to either policy incompleteness or to various forms of "irrationality" (or "behaviorism") in the policy process.

Policy incompleteness arises when policies are based on ad hoc rules that are outcomes not of a deliberate and complete optimisation process, but instead of a much simplified process that may lead to systematic biases in a climate context. One such is case is when discounting of public projects with climate impacts is determined administratively by a common rule for large classes of projects (typically at a high rate), and not aligned with optimality rules relevant for (long-run) climate-related projects. Another case of policy incompleteness is when the returns to public projects are accounted for only over a limited horizon (say, 20 years), or the project is based on no explicit cost-benefit calculation whatsoever.

Considering implications for initial infrastructure design of problems related to categories A-E above, all categories except D are likely to make energy consumption excessive. In cases A-B, this occurs in two complementary ways: through excessive energy intensity of the initial infrastructure; and through excessive ex post "business-as-usual" operation of the investment in period 2. Under case C, initial energy intensity is excessive, while "business-as-usual" operation is here too infrequent (as the initially excessive energy intensity makes it excessively likely that the infrastructure is later retrofitted or abandoned). On balance expected energy use is still typically excessive. The fourth category (D) is different in that it tends to make energy intensity of the initial investment *too low*. An efficient policy would in this case lead to lower total infrastructure cost associated for given energy intensity (as the optimal R&D investment would reduce retrofit costs), and this cost reduction would make an initial energy intensity increase attractive. The infrastructure will however be retrofitted in more cases; this constitutes the main effect for energy consumption, which is reduced in response. In the fifth case, E, the inefficiency could go either way. In one case, excessive optimism over future retrofit possibilities (as reflected in sentiments such as, "future technology will solve everything") could here trigger excessive fossil-fuel intensity in period 1, and overall excessive fuel consumption and emissions.

The second main objective of the paper was to study impacts of the option for later retrofit of an established infrastructure on total expected costs, expected energy (including environmental) costs, and initial energy intensity of infrastructure. These are topics of sections 4 and 6, and of the simulations in section 7. These simulations indicate that some fraction of expected future energy use related to infrastructure can always be avoided by optimally exercising either the retrofit or close-down option at a later stage, given that exercising such options is ex ante optimal. The simulations also indicate that this fraction might be large, under arguably plausible assumptions. In certain parametric cases more than half of the (ex ante expected) potential energy consumption is avoided through optimally exercising the retrofit alternative ex post. Expected total (energy plus retrofit) costs are then reduced, in some cases substantially. These simulations however have limitations. First, they are based on ex ante distributions of energy and retrofit costs that are both log-normal and known, and these two distributions are assumed to be independent. When costs are instead positively correlated, a smaller fraction of the overall expected costs can be avoided by ex post exploitation of low-cost retrofit options. Positive correlation of energy and retrofit costs is here perhaps as likely as cost independence (or more so), as the different cost items that apply to the energy sector may easily co-vary (possibly, with the costs of the whole range of alternative energy technologies, including both renewable energies and pure retrofit technologies co-varying, and fossil-fuel energy costs co-varying with renewables costs).[31] This issue however awaits further analysis, and will be pursued in extensions of this work.

Note that with a long expected time from infrastructure investment to the availability of a relevant alternative option (retrofit, or closedown), and the decision maker discounts heavily (excessively), the options will also be discounted too heavily and given too little weight in the infrastructure decision problem. This factor would then serve as a partial counterweight to those emphasized here, that tend to reduce energy (and climate) cost below socially efficient levels, and lead to too energy intensive infrastructure choices.

Some limitations of our analysis must be pointed out. One is our assumption of only two periods, "the present", and "the future", allowing for only one decision point beyond that of infrastructure investment (at the start of period 2). Our choice of assumptions here was guided by a concern for generality in the distributional assumptions, while still permitting a tractable analysis. An extension of the current framework to three or more periods would make the analysis far less tractable, but should still be pursued in follow-up work. Relevant costs (of

[31] Positive correlation between renewables costs and fossil-fuel costs may in turn follow from a variety of factors. These include downward pressure on fossil fuel market prices from higher levels of renewables production; that less fossil fuels will actually be extracted by any given time frame thus moving less up along the marginal extraction cost curve for such fuels; and that efficiency in fossil fuel extraction may be positively correlated with renewable energy production costs.

energy, emissions, and related to the retrofit technology), as well as benefits (the current utility value of the infrastructure technology), all in reality evolve continuously through time making the two-period framework less accurate. An obvious development would then be to assume that retrofits could be carried out at several points of time; and with separate developments for energy and retrofit costs. Such extensions are in fact considered in a companion paper to the current one, Framstad and Strand (2009), where energy and environmental costs evolve continuously; in other respects however assumptions are here much simpler, in particular, a fixed retrofit cost is assumed.[32]

Another extension is to consider different types of (partial) retrofit, where neither fossil energy use nor carbon emissions are reduced to zero. Such alternatives are typically less costly than full retrofit, and may be more efficient in some cases. Another feature that may need amendment is our assumption that, upon a possible infrastructure closedown, "nothing happens". A more satisfactory analysis would involve a new infrastructure taking over the flow of services lost by closedown; which would require specific assumptions about costs (of new, future, infrastructure investment) and benefits (flowing from the new, instead of the old, infrastructure). We seek to pursue such extensions in future work.

[32] A basic result here is that continuous development of costs produces an "option value" of waiting which serves to delay the retrofit decision; this lowers expected total costs but increases environmental costs.

References

Arthur, Brian (1983), Competing Technologies, Increasing Returns, and Lock-In by Historical Events. *Economic Journal*, 99, 116-131.

Barrett, Scott (2006), Climate Treaties and Breakthrough Technologies. *American Economic Review*, 96 (2), 22-25.

Barrett, Scott (2009), The Coming Global Climate-Technology Revolution. *Journal of Economic Perspectives*, 23 (2), 53-73.

Brueckner, Jan K. (2000), *Economics of Cities – Theoretical Perspectives*. Cambridge, UK: Cambridge University Press.

Buonanno, P., Carraro, C. and Galeotti, M. (2003), Endogenous Induced Technical Change and the Costs of Kyoto. *Resource and Energy Economics*, 25, 11-34.

David, Paul A. (1992), Path Dependence and Predictability in Dynamic Systems With Local Network Externalities: A Paradigm for Historical Economics. In D. Foray and C. Freeman (eds.): *Technology and the Wealth of Nations*, pp 208-231. London: Pinter.

Framstad, Nils Christian and Strand, Jon (2009), Inertia in Infrastructure Investment With "Retrofit": A Continuous-Time Option Value Approach. In process, DEC-EE, World Bank.

Geweke, John (2001), A Note on Some Limitations of CRRA Utility. *Economics Letters*, 71, 341-345.

Glaeser, Edward L. and Kahn, Matthew E. (2008), The Greenness of Cities: Carbon Dioxide Emissions and Urban Development. NBER Working Paper, no 14238.

Golombek, Rolf and Hoel, Michael O. (2005), Climate Policy Under Technology Spillover. *Environmental and Resource Economics*, 31, 201-227.

Golombek, Rolf and Hoel, Michael O. (2006) "Second-Best Climate Agreements and Technology Policy", *Advances in Economic Analysis & Policy*, 6, No. 1, Article 1.

Goulder, L. H. and Mathai, K. (2000), Optimal CO_2 Abatement in the Presence of Induced Technical Change. *Journal of Environmental Economics and Management*, 39, 1-38.

Goulder, L. H. and Schneider, S. H. (1999), Induced Technological Change and the Attractiveness of CO_2 Abatement Policies. *Resource and Energy Economics*, 21, 211-253.

Greaker, Mads and Pade, Lise-Lotte (2008), Optimal CO_2 Abatement and Technological Change. Discussion Paper no 548, Research Department, Statistics Norway.

Gusdorf, François and Hallegatte, Stéphane (2007a), Behaviors and Housing Inertia are Key Factors in Determining the Consequences of a Shock in Transportation Costs. *Energy Policy*, 35, 3483-3495.

Gusdorf, François and Hallegatte, Stéphane (2007b), Compact or Spread-Out Cities: Urban Planning, Taxation and the Vulnerability to Transportation Shocks. *Energy Policy*, 35, 4826-4838.

Ha-Duong, M., Grubb, Michael, and Hourcade, Jean-Charles (1997), Influence of Socioeconomic Inertia and Uncertainty on Optimal CO_2-emission Abatement. *Nature*, 390, 270-273.

Ha-Duong, M. (1998), Quasi-Option Value and Climate Policy Choices. *Energy Economics*, 20, 599-6210.

Hourcade, Jean-Charles and Crassous, Renaud (2008), Low-Carbon Societies: A Challenging Transition for an Attractive Future. *Climate Policy*, 8, 6-0-612.

Jaccard, M. (1997), Heterogeneous Capital Stock and Decarbonating the Atmosphere. Does Delay make Cents? Working Paper, Simon Fraser University.

Jaccard, M. and Rivers, N. (2007), Heterogeneous Captal Stocks and the Optimal Timing for CO_2 Abatement. *Resource and Energy Economics*, 29, 1-16.

Lecocq, Franck, Hourcade, Jean-Charles and Ha-Duong, M. (1998), Decision Making Under Uncertainty and Inertia Constraints: Sectoral Implications of the When Flexibility. *Energy Economics*, 20, 539-555.

Leibowitz, S. and Margulis, S. (1995), Path Dependency, Lock-In and History. *Journal of Law, Economics and Organization*, 11, 205-226.

Lovins, Amory B. (1977), *Soft Energy Paths: Toward a Durable Peace*. Harmondsworth, UK: Penguin Books.

Shalizi, Zmarak and Lecocq, Franck (2009), Economics of Targeted Mitigation Programs in Sectors with Long-Lived Capital Stock. World Bank Working Paper, WPS 5063. Background paper for the World Development Report 2010.

Schuster, Eugene F. (1984), Classification of Probability Laws by Tail Behavior. *Journal of the American Statistical Association*, 79 no 388, 936-939.

Strachan, N., Foxon, T., and Fujino, J. (2008), Modelling Long-Term Scenarios for Low-Carbon Societies. *Climate Policy*, 8 (supplement).

Strachan, N., Pye, S., and Hughes, N. (2008b), The Role of International Drivers on UK Scenarios of a Low-Carbon Soceity. Climate Policy, 8 (supplement), S125-S139.

Strand, Jon (2009), Inertia in Infrastructure Development: Some Analytical Aspects, and Reasons for Inefficient Infrastructure Choices. Unpublished, DECRG, World Bank.

Weitzman, Martin L. (2009), On Modeling and Interpreting the Economics of Catastrophic Climate Change. *Review of Economics and Statistics*, 91, 1-19.

Wigley, T. M. L., Richels, Richard and Edmonds, Jae (1996), Economic and Environmental Choices in the Stabilization of Atmospheric CO_2 Concentrations. *Nature*, 379, 240-243.

World Bank (2003), World Development Report 2003: Sustainable Development in a Dynamic World. Transforming Institutions, Growth, and Quality of Life. Washington DC: The World Bank.

World Bank (2009), World Development Report 2010: Development and Climate Change. Washington DC: The World Bank.